

WHAT IS CLAIMED IS:

1. A system for storing and integrating data entries into a biological data warehouse comprising: a loader module which receives data in a transitional format, converts the transitional format into formatted data, and stores the formatted data in a data warehouse.

5 2. The system of Claim 1, wherein the transitional format comprises a markup language used to represent the data entries.

3. The system of Claim 2, wherein the markup language transforms said data entries into an application and platform-independent form.

10 4. The system of Claim 3, wherein the markup language comprises extensible markup language definitions.

5. The system of Claim 1, wherein the transitional format is converted into a database-compatible language.

15 6. The system of Claim 5, wherein the database-compatible language comprises SQL statements.

7. The system of Claim 1, further comprising a graph generator for generating a data warehouse graph.

20 8. The system of Claim 7, wherein said data warehouse graph is used to represent the schema of said data warehouse, wherein said data entries may be processed in a logical order.

9. The system of Claim 1, further comprising a data verifier for comparing said data entries with data present in said data warehouse.

10. The system of Claim 9, wherein said data verifier is configured to populate incomplete data entries by retrieving the missing information from the data warehouse.

25 11. The system of Claim 1, further comprising a key generator wherein primary and foreign database keys are created within said data warehouse.

12. The system of Claim 1, further comprising a file splitter for splitting large data files to facilitate easier loading of complex data files.

30 13. A method for storing and integrating biological data into a biological data warehouse comprising the steps of:

- a. receiving data in a transitional format;

b. converting said data into a database-compatible language; and
c. storing said database compatible language in a biological data warehouse.

14. The method of claim 13, including the further step of integrating mapping
5 information into the formatted data.

15. The method of claim 14, wherein said mapping information is a mapping file.

16. The method of claim 14, wherein said mapping information is embedded within said transitional formatted data.

10 17. The method of claim 13, wherein said transitional format comprises extensible markup language definitions.

18. The method of claim 13, wherein said database-compatible language comprises SQL statements.

15 19. A system of loading information into a database comprising:
a. translating means for converting data from a transitional format into a database-compatible language;
b. mapping means for corresponding said data with data present in said database; and
c. loading means for storing data into said database.

20